DOCUMENT RESUME

ED 386 023 FL 023 060

AUTHOR Kubota, Mikio
TITLE The Role of Negative Feedback on the Acquisition of the English Dative Alternation by Japanese College Students of EFL.
PUB DATE 94
NOTE 38p.
PUB TYPE Journal Articles (080)
JOURNAL CIT Institute for Research in Language Teaching Bulletin; v8 p1-36 1994

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Classroom Communication; Classroom Research; Classroom Techniques; College Students; Comparative Analysis; *English (Second Language); *Error Correction; *Feedback; Foreign Countries; *Grammar; Higher Education; Instructional Effectiveness; Language Research; Second Language Instruction

ABSTRACT

This study investigated what forms of feedback, if any, helped students of English as a Foreign Language (EFL) learn grammar. Subjects, 100 Japanese college students of EFL, were given two types of tests on English dative ("to/for") alternation, then divided into five groups, according to the type of feedback they received: (1) group A was given information concerning the grammatical rules, as explicit negative feedback; (2) group B was told their answers were incorrect; (3) group C was given correct answers; (4) group D was asked if the answer was right; and (5) group Z, the comparison group, was given no feedback. Results indicate that: group C (modeling and implicit negative feedback) outperformed the comparison group; group A (explicit metalinguistic information) and group C outperformed groups B (explicit utterance rejection) and D (indirect metalinguistic feedback); group B (among B,C, and D) experienced the least effective learning; all experimental groups did better in the first post-test than in the pretest; and no experimental group could use negative feedback to extract linguistic generalizations. It is concluded that all four types of negative feedback had a temporary influence on the learners' linguistic knowledge, and that providing negative feedback in the native language (Japanese) was effective. It is also suggested that teaching rules explicitly and modeling with implicit negative feedback are effective in allowing learners to reformulate their linguistic knowledge. (MSE)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED 386 023

The Role of Negative Feedback on the Acquisition of the English Dative Alternation by Japanese College Students of EFL

紀要第8号抜刷

1994 (Oct 1)

The IRLT Bulletin, Vol 8

PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Mikio Kubota

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

BEST COPY AVAILABLE

財団法人　語学教育研究所

The Institute for Research in Language Teaching

Tokyo, Japan 2

FL023060

The Role of Negative Feedback on the Acquisition of the English Dative Alternation by Japanese College Students of EFL*

Mikio KUBOTA

ABSTRACT

When students produce errors, which kinds of negative feedback teachers provide is one of the important teaching decisions in every class. The purpose of this classroom research is to determine whether negative feedback will help students learn grammar. The subjects, who were 100 Japanese college students of EFL (English as a Foreign Language), were given two kinds of tests on the English dative alternation in each session (Pre-test session, Post-test 1 session, and Post-test 2 session), and they were divided into five groups according to the type of feedback they received. Group A received information concerning the grammatical rules as explicit negative feedback, Group B was told the answers were incorrect, Group C was given correct answers, and Group D was asked if the answer was right. The comparison group (Group Z) received no feedback.

The findings resulting from this classroom research indicate:

(1) Group C (modeling and implicit negative feedback) outperformed the comparison group.

(2) Group A (explicit metalinguistic information) and Group C outperformed Group B (explicit utterance rejection) and Group D (indirect metalinguistic feedback).

(3) Group B among Groups B, C, and D experienced the least effective learning.

(4) All the experimental groups did better in the Post-test 1 than in the

Pre-test.

(5) No experimental groups could use negative feedback to extract linguistic generalizations.

The pedagogical implications are that all four types of negative feedback seem to have a temporary influence on the learners' linguistic knowledge, and that providing negative feedback in Japanese (the native language of the subjects) is effective. Furthermore, it should be important for teachers to remember that teaching rules explicitly and modeling with implicit negative feedback work effectively to let learners reformulate their linguistic knowledge.

1. INTRODUCTION

The kind of feedback teachers should give is an important part of language teaching repertoires, since every lesson inherently includes errors made by students. Besides, from a theoretical point of view, feedback studies do affect inductive learning theory. The mechanisms of induction, which are stimulated by the teacher's feedback, may help explain the development of knowledge of linguistic structures (see Carroll and Swain 1993 in more detail). Therefore, the study of 'negative feedback'-(information to the learner that the learner's production was inappropriate in some way, possibly nonfelicitous, possibly ungrammatical, possibly difficult to parse, etc. Schachter 1993: 182) provides theoretical and practical implications in language teaching methodology for teachers as well as researchers.

First language acquisition research reveals that parents give very little information to children regarding structural properties of language, and that young children do not consciously attend to or understand negative feedback (Brown and Hanlon 1970). Young children cannot interpret negative feedback (Carroll and Swain 1993: 359), because of the limitation of cognitive capacities. There is also anecdotal evidence that children ignore negative feedback (Maratsos 1986). However, the assertion

that negative feedback is not applicable to children has been questioned recently (see White 1989).

In second language acquisition (SLA), naturalistic interaction between native and non-native speakers shows that non-native errors go largely uncorrected (Chun et al. 1982, Day et al. 1984). Classroom research reveals that error correction is not often or systematically available to L2 (second language) learners. Some studies show that teachers did not correct all errors (Allwright 1975, Fanselow 1977, Nystrom 1983), while other studies (Chaudron 1977b/1986, Courchêne 1980, Kubota 1991) show that experienced teachers corrected approximately two-thirds of linguistic errors.

The effect of feedback on developing grammatical competence in the classroom has been carefully investigated. Chaudron (1977a) and Kubota (1991) found that teachers' reduced repetitions with emphasis on the key word (e.g., stress and question intonation) were more strongly correlated to success in revising the original utterance than merely simple or expanded repetitions without emphasis. In addition, Kubota (1991) discovered that teachers' repetitions without change of error (i.e., without changing the error) resulted in success in modification more frequently than repetitions with it.

Tomasello and Herron (1988, 1989), studying American college students learning French as a foreign language, compared two methods for correcting overgeneralization errors and language transfer errors, that is, the "garden path" treatment (learners were led to produce the error, and it was then corrected on the spot by the teacher) and the other treatment (learners were simply taught the exception as an exception). The garden path treatment group performed better than the other treatment group in both immediate and delayed post-tests. This result strongly indicates that on-the-spot error correction works well. Lightbown and Spada (1990) studied the effects of form-focused instruction in ESL (English as a second language) programs that were primarily communi-

cative. These findings suggest that accuracy, fluency, and overall communicative skills are probably best developed through instruction that is primarily meaning-based but in which guidance is provided through timely form-focus activities and correction in context (Lightbown and Spada 1990:443).

Classroom teachers provide a variety of negative feedback for students. Schachter (1984) proposed the following types of feedback: explicit negative feedback and implicit negative feedback (i.e., confirmation check, clarification request, and failure to understand). Carroll et al. (1992) specifically examined the effects of explicit feedback on the learning of French morphological generalizations. Differences were found in favor of experimental groups in a feedback session, but comparisons of guessing responses between experimental and comparison groups showed no evidence of learned generalizations. They also found feedback could help adult second language learners learn individual words. Carroll and Swain (1993) examined the relative effects of various types of negative feedback on the acquisition of the English dative alternation[1] (dative construction or double object construction) by 100 adult Spanish-speaking ESL learners. The results reveal significant differences between all of the feedback groups and the comparison (no feedback) group. Most significantly, the group receiving explicit metalinguistic information regarding the generalizations outperformed the other groups. Their study suggests that both explicit and implicit negative feedback can help learners learn specific and abstract linguistic generalizations.

There has been no research conducted on this topic in EFL specific situations, to the best of this researcher's knowledge. In the present classroom research, I will examine the relative effects of explicit and implicit negative feedback on the acquisition of the English dative alternation by Japanese EFL learners. In order to make a comparison with the results of Carroll and Swain (1993), this is a replication study in most respects.[2]

2. THE STUDY

2.1. RESEARCH QUESTIONS

The main purpose of this research is to investigate whether negative feedback can contribute to the learning of grammar. The following three research questions are thus proposed:

(1) What kinds of negative feedback will lead to the reformulations of subjects' grammatical knowledge in the feedback items?

(2) What kinds of negative feedback will help subjects extract linguistic generalizations in the guessing items from the feedback?

(3) Do the experimental groups perform better after receiving negative feedback?

2.2. PROCEDURES

2.2.1. Subjects

A total of 100 Japanese junior college students in 5 classes participated in this experiment. They had studied EFL in instructional settings for six or seven years. They had already studied the basic usages of the dative alternation while they were in senior high school.

2.2.2. Test Items

The English dative alternation was the target of the syntactic structure in this research, because most Japanese learners of EFL seem to find it very difficult to learn this complex structure. There were two syntactic structures: (1) NP+VP+NP+*to*/*for* NP and (2) NP+VP+NP+NP. In this study, verbs were used with either the preposition *to* or *for*. Not all verbs alternated, and NPs were all lexical.

Two kinds of tests, Test (A) and Test (B) were administered in each session. There were three sessions in this experiment: Pre-test session, Post-test 1 session, and Post-test 2 session. In Test (A) the subjects were required to write out the sentence with two objects if they believed it was grammatically correct. In Test (B) they were asked to write the alternat-

ing form with either *to* or *for*.

During three sessions, the same test items were used in this study. The test items used in the Pre-test were merely rearranged in the Post-tests. And the Post-test 2 was the same test as the Post-test 1.

2.2.3. Research Design

[Step 1] All the subjects in the 5 classes were given the 20-minute Pre-test (see Appendix).

[Step 2] The 5 classes were randomly assigned to the following groups according to the type of feedback the subjects received when they erred. In each class, 20 subjects were chosen randomly.

Group A (20 subjects)—explicit metalinguistic information
Group B (20 subjects)—explicit utterance rejection
Group C (20 subjects)—modeling and implicit negative feedback
Group D (20 subjects)—indirect metalinguistic feedback
Group Z (20 subjects)—no treatment

This grouping followed that of Carroll and Swain (1993) to compare the results of this study with their results. One Japanese teacher of EFL in both experimental and comparison groups was selected in order to control variables (feedback time on task, feedback contents), and he provided feedback in Japanese.

Each subject in the experimental groups was given feedback individually, on a one-to-one basis with the teacher, during the same class period as the Pre-test.

Subjects in Group A were told they were wrong whenever they made an error, and they were given an explicit explanation concerning the dative alternation rule as follows:

Test (A)—if a verb can appear in the syntactic frame V NP *to*/*for* NP, and it expresses transfer of possession in addition to the movement of the theme toward a goal, then it can appear in the syntactic frame V NP NP (Carroll and Swain 1993:363).

Test (B)—the preposition *to* is used in the case of verbs having a directional role, while *for* is used in the case of verbs having a benefactive role.

Group B subjects were simply told they were wrong when they produced an error.

Group C subjects were given a reformulated correct response when they made an error. The negative feedback is implicit, since the subjects were not directly told their response was incorrect.

Group D was asked if the subjects were sure that their response was correct when they made an error. They were not given a model of correct responses.

Group Z, the comparison group received no feedback.

Feedback was provided only in the first-half items of Test (A) and Test (B) in the Pre-test, respectively. The first-half items were called the 'feedback' items (No.1-No.10). On the contrary, in the second-half items, no subjects received feedback. The second-half items were regarded as the 'guessing' items (No.11-No.20). They had to infer the correct responses from the feedback they received in the feedback items.

[Step 3] All the subjects were given the 20-minute Post-test 1 (see Appendix) during the same class period as the Pre-test.

[Step 4] After one month passed, all the subjects were given the 20-minute Post-test 2 (see Appendix).

2.3. HYPOTHESES

Hypotheses 0-4 are concerned with Research Questions (1) and (2), while Hypotheses 5 and 6 are related to Research Question (3).

H0: There would be no statistically significant difference in accuracy of responses between the experimental groups and the comparison group.

That is, the test scores would result in no difference between the experimental groups and the comparison group. If the null hypothesis is incor-

rect, the alternative hypothesis is stated as follows:

H1: The experimental groups would outperform the comparison group in accuracy of responses.

It is assumed that negative feedback would be beneficial in inducing a positive learning effect. Carroll et al. (1992), examining the effects of explicit feedback on the learning of French morphological generalizations, found that experimental groups which received corrective feedback did significantly better than comparison groups which received no feedback, in the feedback sessions. The same result was obtained from Carroll and Swain (1993). Thus, these studies led to the formation of Hypothesis 1.

It is predicted that not all treatments are beneficial to learners. One of the following three hypotheses (H2~H4) would be thus supported.

H2: The group receiving explicit metalinguistic information about generalizations would perform better than the other groups.

This hypothesis means that explicit information about the grammatical rule is the most effective to trigger the learning of the rule and of grammatical generalizations. The most informative type of feedback might consist of a detailed, complete, and accurate grammatical description (Carroll and Swain 1993:362). Therefore, Group A might be the most effective treatment group.

H3: The group receiving the overt model of the desired form would do significantly better than the other groups.

Giving the correct response to learners may help them 'notice the gap' (Schmidt and Frota 1986) between the target norm and the wrong response and then learn the rule. The model of the form (Group C) would be the most helpful in learning the rule.

H4: The groups receiving implicit feedback would perform significantly better than the groups receiving explicit feedback.

If the quantity of information processing required to interpret feedback (Craik and Tulving 1975) is important, the subjects in Groups C and D

could surmise the generalizations more effectively than those in Groups A and B, which received explicit feedback. Implicit information would give sufficient opportunities and time for learners to process and interpret feedback. However, it should be stated that the source of such an error is rather difficult for learners to locate in implicit feedback, because implicit forms of feedback fail to overtly indicate the source of the error.[3]

H5: There would be no statistically significant difference in accuracy of responses among sessions in the experimental groups and the comparison group.

This is also a null hypothesis, because no theory or previous research explains the difference. The following alternative hypothesis will be tested in case the null hypothesis is rejected.

H6: The experimental groups would respond significantly better in the Post-tests 1 and 2 than in the Pre-test.

It is hypothesized that there would exist a positive learning effect longitudinally (after one month), owing to the provision of negative feedback.

2.4. Data Analysis

A .05 level of significance ($\alpha = .05$) was selected. A two-way analysis of variance (ANOVA) with repeated-measures design was employed to analyze the means among the groups.

3. RESULTS

3.1. Test (A)

3.1.1. Feedback Items—Test (A)

The full mark in the feedback items of Test (A) was 10 points. Table 1 shows means and standard deviations by each group and session for the feedback items in Test (A). The means of correct responses for the feedback items in Test (A) are displayed in Figure 1.

Table 2 indicates the results of the two-way repeated-measures

ANOVA for the feedback items in Test (A). The results show that group differences and session differences were statistically significant ($p < .05$). However, the group by session interaction was not statistically significant. This indicates that the main effect for group and the main effect for session were obtained, so that the group and the session influenced the results independently.

Multiple comparisons (using LSD: least square differences) were made in order to determine which groups significantly differed from each other. The results of between-group comparisons are shown in Tables 3, 4, and 5. In the Pre-test, Group A did better than Groups B, D, and Z.

Table 1 : Means and standard deviations by group and session for the feedback items in Test (A)

Group	n	Mean	Standard Deviation
Pre-test			
A	20	7.60	1.24
B	20	6.95	1.24
C	20	7.20	1.21
D	20	6.85	1.49
Z	20	6.70	1.45
Post-test 1			
A	20	8.90	1.61
B	20	7.85	1.19
C	20	8.85	1.11
D	20	8.60	1.43
Z	20	7.65	1.19
Post-test 2			
A	20	8.40	1.32
B	20	6.90	1.51
C	20	8.20	1.12
D	20	7.50	1.47
Z	20	7.50	1.32

Figure 1 : Means of correct responses for the feedback items in Test (A)

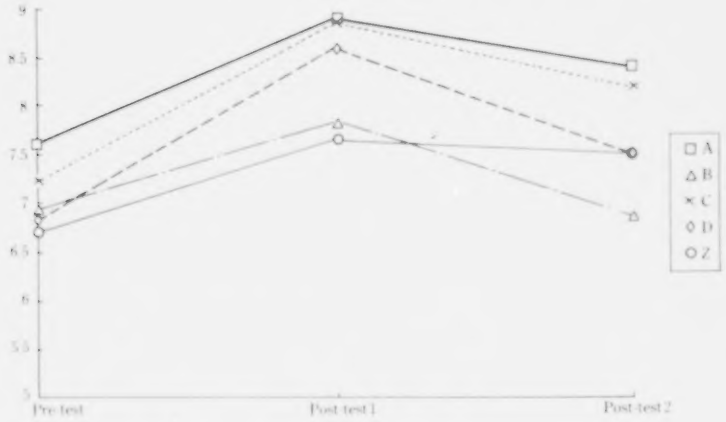

Table 2 : Results of two-way repeated-measures ANOVA for the feedback items in Test (A)

Source	SS	df	MS	F	p
Between subjects	495.00	99			
Groups	80.70	4	20.18	4.63	<.05
Subjects within groups	414.30	95	4.36		
Within subjects	364.69	200			
Session	83.29	2	41.65	29.39	<.05
Groups by session	12.18	8	2.03	1.43	ns
Residual	269.22	190	1.42		

$F_{4,95}$ (.05) = 2.45
$F_{2,190}$ (.05) = 3.00
$F_{8,190}$ (.05) = 1.94

so that these group pairings (A > B, D, Z) were not included in the analysis. Any comparative effects due to treatment are not related to prior knowledge or language ability of any one group, with the exception of these group pairings.

Table 3 : Between-group comparisons of means for the feedback items of Pre-test in Test (A)

Mean	Group	Z	D	B	C
6.70	Z				
6.85	D				
6.95	B				
7.20	C				
7.60	A	*	*	*	

* $p<.05$

Table 4 : Between-group comparisons of means for the feedback items of Post-test 1 in Test (A)

Mean	Group	Z	B	D	C
7.65	Z				
7.85	B				
8.60	D	*	*		
8.85	C	*	*		
8.90	A	(*)	(*)		

* $p<.05$
(*) : excluded in the analysis

Table 5 : Between-group comparisons of means for the feedback items of Post-test 2 in Test (A)

Mean	Group	B	Z	D	C
6.90	B				
7.50	Z	*			
7.50	D	*			
8.20	C	*	*	*	
8.40	A	(*)	(*)	(*)	

* $p<.05$
(*) : excluded in the analysis

In the Post-test 1, the experimental groups (Groups C and D) significantly outperformed the comparison group, whereas in the Post-test 2 Group C did significantly better than the comparison group. Therefore, the null hypothesis (H0) was rejected, thereby Hypothesis 1 was supported.

Among the experimental groups, Groups C and D received significantly higher points than Group B in the Post-test 1. Group C outperformed Group D, which outperformed Group B in the Post-test 2. Thus, the data

Table 6 : Between-session comparisons for means of the feedback items in Test (A)

Mean	Session	Pre	Post 1	Post 2
Group A				
7.60	Pre			
8.90	Post 1	*		
8.40	Post 2	*		
Group B				
6.95	Pre			
7.85	Post 1	*		
6.90	Post 2		*	
Group C				
7.20	Pre			
8.85	Post 1	*		
8.20	Post 2	*		
Group D				
6.85	Pre			
8.60	Post 1	*		
7.50	Post 2	*	*	
Group Z				
6.70	Pre			
7.65	Post 1	*		
7.50	Post 2	*		

* p< .05

reveal the following orderings:

[Post-test 1] C , D > B

[Post-test 2] C > D > B

Since neither Group A nor Group C received the highest scores, Hypotheses 2 and 3 were rejected. Furthermore, Hypothesis 4 was not supported, however Groups C and D performed significantly better than Group B.

The results of between-session comparisons of means for the feedback items in Test (A) are shown in Table 6. Table 6 indicates which sessions were significantly different from each other. Groups A, C, D, and Z

Table 7 : Means and standard deviations by group and session for the guessing items in Test (A)

Group	n	Mean	Standard Deviation
Pre-test			
A	20	7.50	1.32
B	20	6.95	1.40
C	20	7.40	1.16
D	20	6.65	1.82
Z	20	6.95	1.07
Post-test 1			
A	20	7.15	1.24
B	20	7.10	1.04
C	20	7.10	1.18
D	20	6.70	1.31
Z	20	7.00	1.18
Post-test 2			
A	20	7.25	1.13
B	20	6.60	1.24
C	20	6.60	1.16
D	20	6.65	1.19
Z	20	6.30	1.10

Figure 2 : Means of correct responses for the guessing items in Test (A)

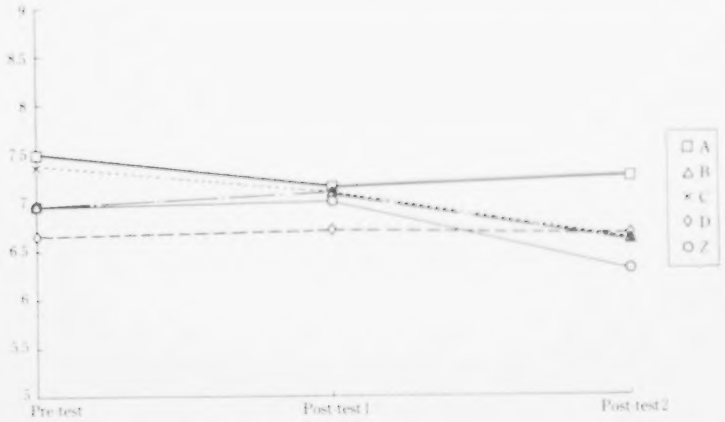

Table 8 : Results of two-way repeated-measures ANOVA for the guessing items in Test (A)

Source	SS	df	MS	F	p
Between subjects	436.73	99			
Groups	26.61	4	6.65	1.54	ns
Subjects within groups	410.12	95	4.32		
Within subjects	209.67	200			
Session	8.67	2	4.34	4.43	<.05
Groups by session	6.27	8	0.78	0.80	ns
Residual	185.73	190	0.98		

$F_{4,95}$ (.05) = 2.45
$F_{2,190}$ (.05) = 3.00
$F_{8,190}$ (.05) = 1.94

performed significantly better in the Post-tests 1 and 2 than in the Pre-test. Besides, Group D performed better in the Post-test 1 than in the Post-test 2. Group B did better in the Post-test 1 than the Pre-test and the Post-test 2. Thus, the null hypothesis (H5) was rejected, thereby

Hypothesis 6 was supported in that all the experimental groups performed better in the Post-tests 1 and 2 than in the Pre-test, with the exception of Group B's Post-test 2.

3.1.2. Guessing Items—Test (A)

The full mark in the guessing items of Test (A) was 10 points. Table 7 shows the respective means and standard deviations by each group and session for the guessing items in Test (A). The means of correct

Table 9: Between-session comparisons of means for the guessing items in Test (A)

Mean	Session	Pre	Post 1	Post 2
Group A				
7.50	Pre			
7.15	Post 1			
7.25	Post 2			
Group B				
6.95	Pre			
7.10	Post 1			
6.60	Post 2			
Group C				
7.40	Pre			
7.10	Post 1			
6.60	Post 2	*		
Group D				
6.65	Pre			
6.70	Post 1			
6.65	Post 2			
Group Z				
6.95	Pre			
7.00	Post 1			
6.30	Post 2		*	

* $p < .05$

responses for the guessing items in Test (A) are illustrated in Figure 2.

Table 8 indicates the results of the two-way repeated-measures ANOVA for the guessing items in Test (A). As indicated in Table 8, session differences (the main effect) were significant, which implies that only sessions influenced the results.

Multiple comparisons were not made to determine which groups were significantly different from each other, since group differences were not significant. Therefore, the null hypothesis (H0) was supported, so that Hypotheses 2-4 did not require testing.

The results of between-session comparisons of means for the guessing

Table 10 : Means and standard deviations by group and session for the feedback items in Test (B)

Group		n	Mean	Standard Deviation
Pre-test				
	A	20	6.60	1.16
	B	20	5.95	1.53
	C	20	6.30	1.35
	D	20	6.15	1.56
	Z	20	5.90	1.51
Post-test 1				
	A	20	7.65	1.24
	B	20	7.05	1.80
	C	20	7.90	1.22
	D	20	6.80	1.66
	Z	20	5.20	2.01
Post-test 2				
	A	20	6.90	1.30
	B	20	6.60	1.43
	C	20	6.45	0.97
	D	20	6.15	1.59
	Z	20	5.35	1.59

items in Test (A) are shown in Table 9. Table 9 indicates which sessions significantly differed from each other. Group C performed significantly worse in the Post-test 2 than in the Pre-test. In addition, Group Z performed worse in the Post-test 2 than in the Post-test 1. Thus, the null

Figure 3 : Means of correct responses for the feedback items in Test (B)

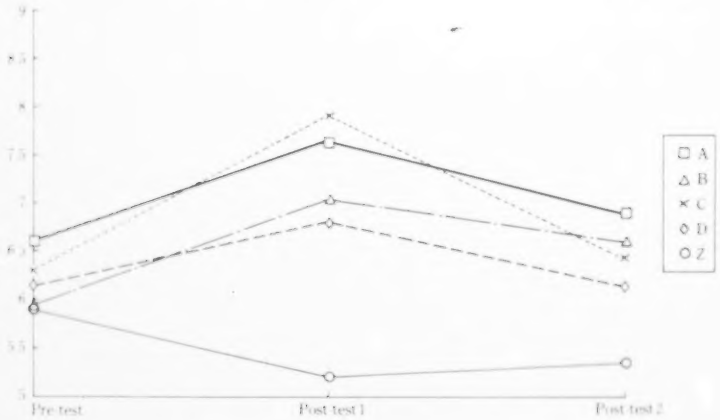

Table 11 : Results of two-way repeated-measures ANOVA for the feedback items in Test (B)

Source	SS	df	MS	F	p
Between subjects	589.24	99			
Groups	124.69	4	31.17	6.37	$<.05$
Subjects within groups	464.55	95	4.89		
Within subjects	347.33	200			
Session	30.98	2	15.49	10.47	$<.05$
Groups by session	35.65	8	4.46	3.01	$<.05$
Residual	280.70	190	1.48		

$F_{4,95}(.05) = 2.45$
$F_{2,190}(.05) = 3.00$
$F_{8,190}(.05) = 1.94$

hypothesis (H5) was rejected, and Hypothesis 6 was supported in that only one experimental group (Group C) performed worse in the Post-test 2 than in the Pre-test.

3.2. Test (B)

3.2.1. Feedback Items—Test (B)

The full mark in the feedback items of Test (B) was 10 points. Table 10 shows means and standard deviations by each group and session for the feedback items in Test (B). The means of correct responses for the feedback items in Test (B) are illustrated in Figure 3.

Table 11 indicates the results of the two-way repeated-measures ANOVA for the feedback items in Test (B). The results show that group differences and session differences were significant ($p<.05$). In addition, the group by session interaction was statistically significant. Since the main effect and the interaction were significant, groups and sessions influenced the results dependently.

Multiple comparisons (using LSD) were made in order to determine which groups significantly differed from each other. The results of between-group comparisons are shown in Tables 12, 13, and 14. In the Pre-test, Group A outperformed Groups B and Z, so that these group pairings (A > B, Z) were not counted in this analysis. In both Post-tests 1 and 2, the experimental groups (Groups B, C, D) significantly outperformed the comparison group. Therefore, the null hypothesis (H0) was rejected, thereby Hypothesis 1 was supported.

Among the experimental groups, in both Post-tests 1 and 2, Group A did better than Group D, and the mean of Group C was significantly higher than means of Groups B and D in the Post-test 1. The data reveal the following orderings:

[Post-test 1] A > D
C > B, D
[Post-test 2] A > D

Table 12 : Between-group comparisons of means for the feedback items in Pre-test of Test (B)

Mean	Group	Z	B	D	C
5.90	Z				
5.95	B				
6.15	D				
6.30	C				
6.60	A	*	*		

* $p < .05$

Table 13 : Between-group comparisons of means for the feedback items in Post-test 1 of Test (B)

Mean	Group	Z	D	B	A
5.20	Z				
6.80	D	*			
7.05	B	*			
7.65	A	(*)	*		
7.90	C	*	*	*	

* $p < .05$
(*) : excluded in the analysis

Table 14 : Between-group comparisons of means for the feedback items in Post-test 2 of Test (B)

Mean	Group	Z	D	C	B
5.35	Z				
6.15	D	*			
6.45	C	*			
6.60	B	*			
6.90	A	(*)	*		

* $p < .05$
(*): excluded in the analysis

Table 15 : Between-session comparisons of means for the feedback items in Test (B)

Mean	Session	Pre	Post 1	Post 2
Group A				
6.60	Pre			
7.65	Post 1	*		
6.90	Post 2		*	
Group B				
5.95	Pre			
7.05	Post 1	*		
6.60	Post 2	*		
Group C				
6.30	Pre			
7.90	Post 1	*		
6.45	Post 2		*	
Group D				
6.15	Pre			
6.80	Post 1	*		
6.15	Post 2		*	
Group Z				
5.90	Pre			
5.20	Post 1	*		
5.35	Post 2			

* $p < .05$

Since the difference between Groups A and C reached no statistical significance', Hypotheses 2 and 3 were rejected. Furthermore, Hypothesis 4 was not supported.

The results of between-session comparisons of means for the feedback items in Test (B) are shown in Table 15. Table 15 indicates which sessions were significantly different from each other. All the experimental groups (Groups A, B, C, D) performed significantly better in the Post-test 1 than in the Pre-test. Furthermore, Groups A, C, and D

performed better in the Post-test 1 than in the Post-test 2. Group B did better in the Post-test 2 than the Pre-test. Thus, the null hypothesis (H5) was rejected, thereby Hypothesis 6 was partly supported in that all the experimental groups performed better in the Post-test 1 than in the Pre-test.

3.2.2. Guessing Items—Test (B)

The full mark in the guessing items of Test (B) was 10 points. Table 16 shows means and standard deviations by each group and session for the guessing items in Test (B). The means of correct responses for the

Table 16 : Means and standard deviations by group and session for the guessing items in Test (B)

Group		n	Mean	Standard Deviation
Pre-test				
	A	20	7.45	1.40
	B	20	6.85	2.22
	C	20	7.70	1.42
	D	20	7.25	1.89
	Z	20	6.85	1.80
Post-test 1				
	A	20	7.10	1.45
	B	20	7.35	1.43
	C	20	7.85	1.11
	D	20	7.00	1.67
	Z	20	7.20	1.25
Post-test 2				
	A	20	7.05	1.50
	B	20	7.55	1.47
	C	20	7.85	1.31
	D	20	7.05	1.53
	Z	20	7.00	1.67

Figure 4 : Means of correct responses for the guessing items in Test (B)

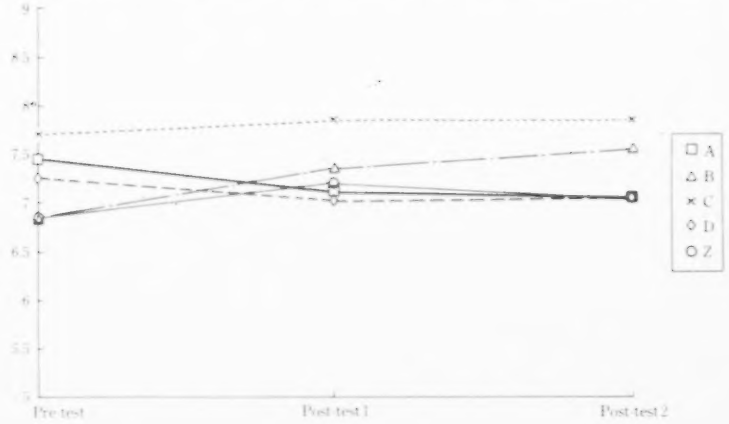

Table 17 : Results of two-way repeated-measures ANOVA for the guessing items in Test (B)

Source	SS	df	MS	F	P
Between subjects	591.33	99			
Groups	40.50	4	10.13	1.75	ns
Subjects within groups	550.83	95	5.80		
Within subjects	330.67	200			
Session	0.38	2	0.19	0.11	ns
Groups by session	9.02	8	1.13	0.67	ns
Residual	321.27	190	1.69		

$F_{4,95}(.05) = 2.45$
$F_{2,190}(.05) = 3.00$
$F_{8,190}(.05) = 1.94$

guessing items in Test (B) are illustrated in Figure 4.

Table 17 indicates the results of the two-way repeated-measures ANOVA for the guessing items in Test (B). As shown in Table 17, group

differences and session differences were not significant in the guessing items of Test (B). Moreover, the group by session interaction was not significant.

Multiple comparisons were not made to determine which groups significantly differed from each other, since group differences were not of consequence. Therefore, the null hypothesis (H0) was supported, so that Hypotheses 2-4 were not tested.

The analysis of between-session comparisons of means for the guessing items in Test (B) was not made, since session differences were not statistically significant. Thus, the null hypothesis (H5) was supported.

4. DISCUSSION

The following hypotheses were tested in this study:

H0: There would be no statistically significant difference in accuracy of responses between the experimental groups (A, B, C, D) and the comparison group (Z).

H1: The experimental groups would outperform the comparison group in accuracy of responses.

H2: The group (A) receiving explicit metalinguistic information about generalizations would perform better than the other groups.

H3: The group (C) receiving the overt model of the desired form would do significantly better than the other groups.

H4: The groups (C, D) receiving implicit feedback would perform significantly better than the groups (A, B) receiving explicit feedback.

H5: There would be no statistically significant difference in accuracy of responses among sessions in the experimental groups and the comparison group.

H6: The experimental groups (A, B, C, D) would respond significantly better in the Post-tests 1 and 2 than in the Pre-test.

4.1. Feedback Items

H0 H1: In Test (A), which required the subjects to write the sentence with two objects, H1 was supported in that Group C outperformed Group Z (comparison group) in the Post-tests 1 and 2. Although Group D did better than Group Z in the Post-test 1, the same result was not obtained in the Post-test 2. This clearly shows that a learning effect in Group D disappeared during the three tests. In Test (B), which asked the subjects to write the sentence with *to* or *for*, Groups B, C, and D performed better than the comparison group in the Post-tests 1 and 2. This result shows that the learning effect was maintained by the provision of negative feedback (except Group A) in the feedback items of Test (B), in which the teacher gave negative feedback to the subjects in Japanese. Therefore, in light of the results of Tests (A) and (B), it is concluded that Group C outperformed the comparison group.

H2 H4: In Test (A), the results reveal that Group B received the lowest points in the Post-tests 1 and 2, among Groups B, C, and D, thereby negative feedback of explicit utterance rejection used in Group B (the group which was told that the responses were wrong) led to the least effective learning among Groups B, C, and D. Even Group Z performed better than Group B in the Post-test 2. Moreover, the learning effect in Group D did not continue during the three tests; in the Post-test 1 the difference between Groups C and D reached no statistical significance, while in the Post-test 2 Group C outperformed Group D.

In Test (B), Group A outperformed Group D in the Post-tests 1 and 2. The learning effect in Group C dropped seriously; in the Post-test 1 Group C did better than Groups B and D, whereas in the Post-test 2 a statistically significant difference was not obtained among Groups B, C, and D. This trend proves difficult to interpret: the subjects in Group C might have difficulty recalling the correct form provided by the teacher, that is, whether *to* or *for* should be used.

These results warrant the following orderings:

Test (A): C, D > B [Post-test 1]
C > D > B [Post-test 2]
Test (B): A > D [Post-test 1]
C > D, B [Post-test 1]
A > D [Post-test 2]

It can be stated that overall, Groups A and C outperformed Groups B and D, and that Group B experienced the least effective learning among Groups B, C, and D.

H5 H6 In Test (A), all the experimental groups did significantly better in the Post-tests 1 and 2 than in the Pre-test, except Group B's Post-test 2. In Test (B), all the experimental groups performed better in the Post-test 1 than in the Pre-test. Moreover, it should be noted that Group B in Test (A) did not continue its learning effect over a one-month period, whereas Group B in Test (B) did continue it; Group B in Test (B) did better in the Post-test 2 than in the Pre-test. The other experimental groups in Test (B) did not show statistical differences between the Pre-test and the Post-test 2. So, the effect of explicit utterance rejection used in Group B gave contradictory results. This topic deserves further investigation.

It is claimed that all the experimental groups generated significantly better responses in the Post-test 1 than in the Pre-test. This research, thus, provides evidence that negative feedback is effective, at least temporarily—though the subjects did not benefit from negative feedback in the long term. Ellis (1993: 17) discussed what factors will determine whether the effect of grammar instruction persists: (1) whether the target items are frequent in classroom input and output to learners after treatment, (2) what the nature of the linguistic feature itself is—a developmental feature (e.g., questions) or a variational feature (e.g., adverbs), and (3) how the learner perceives the importance of a given grammatical feature. It may be important for teachers to speculate whether learners retain a given grammatical structure in the short term

or in the long term, and what factors determine the result. At any rate, long-term learning durability remains at issue.[4]

It is interesting to note that the subjects of Group Z showed variability regarding their responses: in the feedback items of Test (A), Group Z did significantly better in the Post-test 1 than in the Pre-test (7.65 vs. 6.70), while in the feedback items of Test (B) Group Z performed better in the Pre-test than in the Post-test 1 (5.90 vs. 5.20). This variability may be influenced by test items, test fatigue, and the subjects' instability of linguistic knowledge. Conversely, it is fair to say that in this research, the provision of negative feedback enabled the subjects to reduce the possibility of variability of their responses, since all the experimental groups got significantly higher scores in the Post-test 1 than in the Pre-test in the feedback items of both Test (A) and Test (B), as illustrated in Tables 6 and 15.

4.2. Guessing Items

H0 H4. In Tests (A) and (B), the null hypothesis (H0) was supported, so that there was no statistically significant difference in accuracy of responses between the experimental groups (A, B, C, D) and the comparison group (Z). This clearly indicates that the guessing items failed to obtain significant results, which remains consistent with the results of Carroll et al. (1990, 1992). The data in this study prove that the subjects did not extract the expected linguistic generalizations from the feedback that they were applying to new items. The answer to this may be that this phenomenon is due to "the limited amount of time and exposure to the linguistic system" (Nayak et al. 1990).

H5 H6. In Test (A), only Group C performed worse in the Post-test 2 than in the Pre-test. In Test (B), the null hypothesis (H5) was supported in that there was no statistically significant difference in accuracy of responses among sessions in the experimental groups and the comparison group. Thus, negative feedback did not lead to learning linguisitic

generalizations and did not alter learners' linguistic knowledge, either.

Carroll and Swain (1993) found that the group given an explicit rule about the dative alternation (Group A) performed significantly better than all other groups in both the feedback items and the guessing items. In this research, however, the same result was not obtained in either the feedback items or the guessing items.

5. CONCLUSION

The following major findings and pedagogical implications are presented, based on this classroom research.

The results from the feedback items indicate:

(1) Group C (modeling and implicit negative feedback) outperformed the comparison group.

(2) Groups A (explicit metalinguistic information) and Group C outperformed Group B (explicit utterance rejection) and Group D (indirect metalinguistic feedback).

(3) Group B among Groups B, C, and D experienced the least effective learning.

(4) All the experimental groups did better in the Post-test 1 than in the Pre-test.

All four types of negative feedback seem to have a temporary influence on the learners' linguistic knowledge. It should be pointed out that negative feedback, which proved to be successful, was provided in Japanese, the native language of the subjects. It might also be important for teachers to remember that teaching rules explicitly (Group A) and modeling and giving implicit negative feedback (Group C) are influential to let learners reformulate their linguistic knowledge.

The results from the guessing items indicate:

(5) No experimental groups could use negative feedback to extract linguistic generalizations.

Teachers may keep in mind that the amount of exposure to linguistic

structures is a key factor of induction.

It should be noted that this classroom research derives from small-scale research (i.e., No. of tests given: 2, No. of subjects:100), so that it may be improper to make firm conclusions. This indicates only as a pilot experiment that was first conducted in Japan. In addition, group equating could be incorporated into this type of experiment so that any group pairings would not be excluded in the analysis.

Further research should replicate this experiment, using various other production/reception tasks (e.g., oral production task/grammaticality judgment task) with the same and/or different linguistic items in order to investigate whether learning about the language is useful or not in facilitating language development and what kind of negative feedback is most effective for long-term retention.

NOTES

* This research was partly supported by a grant from the Foundation for Promoting Language Teaching (Gengo Kyoiku Shinkou Zaidan), awarded in 1994.

1. Mazurkewich (1984) found that unmarked structures, or dative prepositional phrase complements were acquired before marked structures, or double-NP complements by ESL learners, stating that this evidence provided support for a theory of markedness. Hawkins (1987) confirmed the result of Mazurkewich (1984) and offered another account of the acquisition process in terms of learning complexity. The syntactic distributional subclass of the verb interacts to produce a multi-staged developmental sequence. L2 learners begin with a distinction between pronominal and lexical datives, which is later refined by the introduction of a distinction between *to*-verbs and *for*-verbs, and later still is refined by the introduction of the distinction between native and nonnative verb forms (Hawkins 1987: 46). The acquisition process is represented as [±pronominal] → [±to] → [±native]. These analyses, though significant, are not within the scope of the present study.

2. This research is unique in that two kinds of tests were given to the subjects and both between-group comparisons and between-session comparisons were made as multiple comparisons in analyzing the data.

3. Kubota (1991) found that explicit corrective feedback was more successful in triggering students' modified correct forms than implicit corrective feedback.

BEST COPY AVAILABLE

4. There have been few studies which discovered a long-term effect for focusing on form (e.g., Lightbown 1991, Spada and Lightbown 1993).

REFERENCES

Allwright, R.L. "Problems in the study of the language teachers' treatment of learner error." In Burt, M and H. Dulay, eds. (1975), 96-109.

Brown, R. and C. Hanlon, "Derivational complexity and order of acquisition in child speech." In Hayes, J.R., ed. (1970), 11-53.

Burt, M and H. Dulay, eds. *On TESOL '75: New Directions in Second Language Learning, Teaching and Bilingual Education*. Washington, DC: TESOL, 1975.

Carroll, S., Y. Roberge, and M. Swain. "Error correction and the acquisition of word formation rules." Paper presented at the 10th Second Language Research Forum, University of Oregon, Eugene, 1990.

———. "The role of feedback in adult second language acquisition: error correction and morphological generalizations." *Applied Linguistics*, 13, 2 (1992), 173-198.

Carroll, S. and M. Swain. "Explicit and implicit negative feedback: an empirical study of the learning of linguistic generalizations." *Studies in Second Language Acquisition*, 15, 3 (1993), 357-386.

Chaudron, C. "A descriptive model of discourse in the corrective treatment of learner's errors." *Language Learning*, 27, 1 (1977a), 29-46.

———. "Teachers' priorities in correcting learners' errors in French immersion classes." *Working Papers on Bilingualism*, 12 (1977b), 22-44. Reprinted in Day, R. R., ed. (1986), 64-84.

Chun, A., R. Day, N. Chenoweth, and S. Luppescu. "Errors, interaction, and correction: a study of native-nonnative conversation." *TESOL Quarterly*, 16, 4 (1982), 537-547.

Courchêne, R. "The error analysis hypothesis, the contrastive analysis hypothesis, and the correction of error in the second language classroom." *TESL Talk*, 11, 2 (1980), 3-13; 11, 3 (1980), 10-29.

Craik, F.I.M. and E. Tulving. "Depth of processing and the retention of words in episodic memory." *Journal of Experimental Psychology: General*, 104 (1975), 268-294.

Day, R.R., ed. *Talking to Learn: Conversation in Second Language Acquisition*. Cambridge, MA: Newbury House, 1986.

Day, R., N. Chenoweth, A. Chun, and S. Luppescu. "Corrective feedback in native-nonnative discourse." *Language Learning*, 34 (1984), 19-45.

Ellis, R. "Grammar teaching and grammar learning." *Temple University Japan Research Studies in TESOL*, 1, 1 (1993), 3-27.

Fanselow, J.F. "The treatment of error in oral work." *Foreign Language Annals*, 10, 5 (1977), 583-593.

Hawkins, R. "Markedness and the acquisition of the English dative alternation by L2 speakers." *Second Language Research*, 3 (1987), 20-55.

Hayes, J.R., ed. *Cognition and Development of Language*. New York: Wiley, 1970.

Kubota, M. "Corrective feedback by experienced Japanese EFL teachers." *The IRLT Bulletin*, 5 (1991), 1-25. Tokyo: The Institute for Research in Language Teaching.

Lightbown, P.M. "What have we here? Some observations on the influence of instruction on L2 learning." In Phillipson, R., E. Kellerman, M. Sharwood Smith, and M. Swain, eds. (1991), 197-212.

Lightbown, P.M. and N. Spada. "Focus-on-form and corrective feedback in communicative language teaching: effects on second language learning." *Studies in Second Language Acquisition*, 12, 4 (1990), 429-448.

Maratsos, M.P. "On the roles in input and tabulation." Paper presented at the 11th annual Boston University Conference on Language Development, Oct. 17-19, 1986.

Muzurkewich, I. "The acquisition of the dative alternation by second language learners and linguistic theory." *Language Learning*, 34, 1 (1984), 91-109.

Nayak, N., N. Hansen, N. Krueger, and B. McLaughlin. "Language-learning strategies in monolingual and multilingual subjects." *Language Learning*, 40, 2 (1990), 221-244.

Nystrom, N.J. "Teacher-student interaction in bilingual classrooms: four approaches to error feedback." In Seliger, H.W. and M.H. Long, eds. (1983), 169-189.

Phillipson, R., E. Kellerman, M. Sharwood Smith, and M. Swain, eds. *Foreign/Second Language Pedagogy Research*. Clevedon, England: Multilingual Matters, 1991.

Rutherford, W.E., ed. *Language Universals and Second Language Acquisition*. Philadelphia: John Benjamins, 1984.

Schachter, J. "A universal input condition." In Rutherford, W.E., ed. (1984), 167-183.

———. "Second language acquisition: perceptions and possibilities." *Second Language Research*, 9, 2 (1993), 173-187.

Schmidt, R.W. and S.N. Frota. "Developing basic conversational ability in a second language: a case study of an adult learner of Portuguese." In Day, R.R., ed. (1986), 237-327.

Seliger, H.W. and M.H. Long, eds. *Classroom-Oriented Research in Second Language Acquisition*. Rowley, MA: Newbury House, 1983.

Spada, N. and P. Lightbown. "Instruction and the development of questions in L2

classrooms." *Studies in Second Language Acquisition*, 15, 2 (1993), 205-224.

Tomasello, M. and C. Herron. "Down the Garden Path: inducing and correcting overgeneralization errors in the foreign language classroom." *Applied Psycholinguistics*, 9, 3 (1988), 237-246.

———. "Feedback for language transfer error: the Garden Path technique." *Studies in Second Language Acquisition*, 11, 4 (1989), 385-395.

White, L. *Universal Grammar and Second Language Acquisition*. Amsterdam: John Benjamins, 1989.

APPENDIX

[Pre-test]

Test (A):

Write out the sentence with two objects, if you believe it is grammatically correct, as in the following (a). If a verb does not allow two objects, write X, as in the following (b).

E.g., (a) I showed a picture to Mary.
I showed Mary a picture.
(b) She introduced her new friend to Ken.
X

1. John gave a pen to his sister.
2. I bought a ring for my friend.
3. Could you open the door for me?
4. He recommended a book to his brother.
5. I wrote a letter to Tom.
6. He left a note for his son.
7. Tom drove a car for Kate.
8. He told a story to Jane.
9. Jim carried a bag for Jane.
10. Will you pass the salt to Tom?
11. John found a job for Paul.
12. My father cooks supper for us.
13. I want to save money for a trip.
14. They paid the fees to the manager.
15. The teacher reported the news to her students.
16. He handed a book to Jerry.
17. I cleaned the room for Harry.
18. He painted a picture for Mike.
19. They offered a job to the man.
20. The father brought a cat for his child.

Test (B):

Write out the sentence with a preposition ***to*** (as in (a)) or ***for*** (as in (b)).

E.g., (a) I showed Mary a picture

I showed a picture ***to*** Mary.

(b) I chose my mother that book.

I chose that book ***for*** my mother.

1. The clerk sold Linda a bag.
2. I ordered everyone beer.
3. I will mail Jane this letter.
4. Tom promised his son a present.
5. John made his sister a bookshelf.
6. She denied her child an apple.
7. Can you lend me 100 yen?
8. Please telephone me your reply.
9. I prepared John a report.
10. Could you get me a drink?
11. She read her child a story.
12. He baked his mother apples.
13. I allowed Mary $20.
14. She caught her friend fish.
15. Your letter caused Mary trouble.
16. He threw Ken a ball.
17. They refused Mike admission.
18. I reserved my son a room.
19. I wished Ken a happy journey.
20. Could you spare me a few minutes?

[Post-test]

(NB: Post-test 2 is the same test as Post-test 1)

Test (A):

Write out the sentence with two objects, if you believe it is grammatically correct, as in the following (a). If a verb does not allow two objects, write X, as in the following (b).

E.g., (a) I showed a picture to Mary.
I showed Mary a picture.

(b) She introduced her new friend to Ken.
X

1. Will you pass the salt to Tom?
2. He painted a picture for Mike.
3. I want to save money for a trip.
4. He left a note for his son.
5. John gave a pen to his sister.
6. He handed a book to Jerry.
7. He told a story to Jane.
8. John found a job for Paul.
9. The teacher reported the news to her students.
10. They offered a job to the man.
11. I bought a ring for my friend.
12. They paid the fees to the manager.
13. Tom drove a car for Jane.
14. The father brought a cat for his child.
15. He recommended a book to his brother.
16. Jim carried a bag for Jane.
17. Could you open the door for me?
18. My father cooks supper for us.
19. I cleaned the room for Harry.
20. I wrote a letter to Tom.

Test (B):

Write out the sentence with a preposition ***to*** (as in (a)) or ***for*** (as in (b)).

E.g., (a) I showed Mary a picture
I showed a picture ***to*** Mary.

(b) I chose my mother that book.
I chose that book ***for*** my mother.

1. He baked his mother apples.
2. Tom promised his son a present.
3. They refused Mike admission.
4. I prepared John a report.
5. She caught her friend fish.
6. She denied her child an apple.
7. Your letter caused Mary trouble.
8. Could you spare me a few minutes?
9. The clerk sold Linda a bag.
10. I allowed Mary $20.
11. She read her child a story.
12. I will mail Jane this letter.
13. I wished Ken a happy journey.
14. Can you lend me 100 yen?
15. I ordered everyone beer.
16. I reserved my son a room.
17. Please telephone me your reply.
18. He threw Ken a ball.
19. John made his sister a bookshelf.
20. Could you get me a drink?

(調布学園女子短期大学)